

# HAB (Harmful Algal Bloom) Detection System

COMP47250 Team Software Project Final Report

Daniel Ilyin  
*School of Computer Science, UCD*  
Dublin, Ireland  
daniel.ilyin@ucdconnect.ie

Sagar Satish Poojary  
*School of Computer Science, UCD*  
Dublin, Ireland  
email address @ucdconnect.ie

Dharmik Arvind Vara  
*School of Computer Science, UCD*  
Dublin, Ireland  
email address @ucdconnect.ie

Karthika Garikapati  
*School of Computer Science, UCD*  
Dublin, Ireland  
email address @ucdconnect.ie

Kruthi Jangam  
*School of Computer Science, UCD*  
Dublin, Ireland  
email address @ucdconnect.ie

Roshan Palem  
*School of Computer Science, UCD*  
Dublin, Ireland  
email address @ucdconnect.ie

**Abstract**—Harmful Algal Blooms (HABs) pose serious risks to marine ecosystems, human health, and coastal economies. Yet most traditional detection methods remain slow, expensive, and limited in how much area they can cover. In this project, we developed a smarter, faster alternative: an AI-powered HAB Detection System that uses satellite imagery and deep learning to predict toxic bloom events more effectively.

At the core of our system is a CNN+LSTM model trained on spatiotemporal data-cubes built from satellite-derived environmental data. The model takes in three key modalities chlorophyll-a concentration, sea surface temperature, and remote sensing reflectance captured across a 10-day window. To make the system accessible to a range of users, we designed a tiered architecture: a lightweight Free Tier using a single modality over 5 days, and more advanced tiers combining richer data for higher accuracy.

Users interact with the system through a simple web interface where they can click on a map to get instant predictions for any location. Our best performing model (Tier 2) achieved 92% accuracy, while the lighter tiers offered reasonable performance with reduced data input and compute needs.

By combining geospatial interactivity, deep learning, and flexible design, this system offers a practical step forward in early HAB detection especially for resource-limited agencies that need timely, reliable insights.

*Index Terms*—

## I. INTRODUCTION

Harmful Algal Blooms (HABs) are showing up more frequently and with greater intensity across coastal waters around the world. These toxic blooms, often caused by nutrient runoff and warming ocean temperatures, can do serious damage to marine life, public health, and local economies. In many cases, they've led to large-scale fish deaths, contaminated drinking water, and forced the shutdown of both recreational beaches and fishing operations. Despite the growing threat, most current methods of detecting HABs still rely on manual sampling and lab analysis which are slow, expensive, and not practical for covering large areas in real time.

Our project set out to build a smarter, more scalable approach using satellite data and deep learning. The aim was to design a system that can predict toxic bloom events from space quickly, accurately, and in a way that's easy for people

to use. It fits into a larger effort to apply AI to real-world environmental problems, where timely action can help prevent serious harm.

The main challenge we faced was figuring out how to model the way key environmental variables shift over both space and time. To solve that, we used a hybrid model combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, trained on spatiotemporal data-cubes built from satellite imagery. These data-cubes cover 10 days of data across three critical variables: chlorophyll-a levels, sea surface temperature, and remote sensing reflectance.

To make the system flexible and widely usable, we also introduced a tiered prediction structure. Users can choose between a lightweight free version which uses minimal data and more advanced tiers that deliver better accuracy using richer inputs. All of this is wrapped into a simple web interface where users can click on a map and get bloom predictions instantly. The result is a system that combines technical depth with real-world accessibility to stay ahead of harmful algal bloom events.

## II. USER STORIES AND SCENARIOS

- A. *Marine Biologists and Environmental Agencies*
- B. *Aquaculture Operators*
- C. *Public Health Officials*

## III. STATE OF THE ART AND RELATED WORK

- A. *Traditional HAB Detection Methods*
- B. *Remote Sensing Approaches*
- C. *Machine Learning in Environmental Monitoring*
- D. *Spatiotemporal Analysis Methods*
- E. *Comparison with Existing Systems*

## IV. IMPLEMENTATION

- A. *System Architecture Overview*
- B. *Data Collection and Preprocessing Pipeline*
  - 1) *Ground Truth Data Sources:*

- 2) *Remote Sensing Data Acquisition:*
- 3) *Data Preprocessing and Quality Control:*
- C. *Datacube Construction*
  - 1) *Spatiotemporal Window Definition:*
  - 2) *Modality Selection and Integration:*
  - 3) *Sparse Data Handling:*
- D. *Machine Learning Architecture*
  - 1) *CNN Feature Extraction:*
  - 2) *Temporal Classification Methods:*
  - 3) *Model Training and Validation:*
- E. *Web-Based Dashboard*
  - 1) *Frontend Implementation:*
  - 2) *Real-time Data Integration:*
  - 3) *Visualization Components:*
- F. *Cloud Infrastructure and Deployment*
  - 1) *AWS Architecture:*
  - 2) *MLOps Pipeline:*
  - 3) *Scalability and Performance Optimization:*

## V. EVALUATION AND RESULTS

- A. *Dataset Description*
- B. *Experimental Setup*
  - 1) *Cross-Validation Strategy:*
  - 2) *Performance Metrics:*
  - 3) *Baseline Comparisons:*
- C. *Detection Performance*
  - 1) *Classification Results:*
  - 2) *Feature Importance Analysis:*
  - 3) *Comparison with Traditional Methods:*
- D. *Prediction Performance*
  - 1) *Temporal Forecasting Results:*
  - 2) *Prediction Horizon Analysis:*
- E. *System Performance Evaluation*
  - 1) *Scalability Testing:*
  - 2) *Response Time Analysis:*
  - 3) *Resource Utilization:*
- F. *User Evaluation*
  - 1) *User Study Design:*
  - 2) *Usability Assessment:*
  - 3) *User Feedback and Insights:*
- G. *Technical Evaluation*
  - 1) *Model Performance Analysis:*
  - 2) *Infrastructure Performance:*
  - 3) *System Reliability:*

## VI. CONCLUSIONS AND FUTURE WORK

- A. *Summary of Achievements*
- B. *Key Technical Contributions*
- C. *Limitations and Challenges*
- D. *Future Directions*
  - 1) *Model Improvements:*
  - 2) *Data Enhancement:*
  - 3) *System Extensions:*
- E. *Impact and Applications*

## ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

## REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.